The Application of COSMIC Data to Global Change Research [1]

Stephen S. Leroy
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Gerald R. North
Department of Meteorology
Texas A&M University
College Station, Texas

2 August 1999

[1] Corresponding author address: Dr. Stephen S. Leroy, Jet Propulsion Laboratory, MS 169-237, 4800 Oak Grove Dr., Pasadena, California 91109 (USA); email: Stephen.S.Leroy@jpl.nasa.gov.

1   Introduction

Carbon dioxide concentration in the atmosphere is rising at about 0.5% per year, almost entirely due to emissions of fossil fuel combustion, and it is expected to cause increases in surface temperature measurements. All of this is completely addressed in the reports of the Intergovernmental Panels on Climate Change (IPCC) (Houghton et al. 1991; Houghton and Meira Filho 1996). The overriding focus of the research summarized in the IPCC reports is to understand the effects greenhouse gas forcing has had in the past and the effects it is likely to have in the future. In this paper we do not choose to outline all of those issues addressed in the IPCC reports but instead outline only climate signal detection and attribution and the statistical testing of climate models. We lay out the mathematics of each keeping in mind how occultation data obtained by COSMIC will apply. A commonly occuring theme will be the importance of testing and improving climate models.

The issue of detecting changes in the climate and attributing those changes to specific "forcings" is a general one which has most frequently been applied to warming of Earth's surface together with anthropogenic greenhouse gas increases. While the "attribution" problem is highly relevant for public policy purposes, its scientific value is just beginning to be put into context: detection and attribution tells more about the quality of climate models than it does about cause and effect in the climate system. Namely, the effects of any specific climate forcing can only be hypothesized by the use of a climate model—global climate models (GCMs) being the most sophisticated—which themselves give highly inconsistent results. In fact, GCMs predict anywhere from a 2 K to a 5 K increase in surface temperature due to a doubling of carbon dioxide concentration. Clearly, the models themselves cannot be satisfactory predictors of climate change yet, so it is the data that we must rely upon to update the models. Hence, commonly it is the climate "sensitivity", or the surface temperature increase caused by a doubling of CO_2, which is sought out in data. Here we lay out the most current versions of climate signal detection, how one would implement occultation data from COSMIC, describe recent results in the literature, and show rigorously this process is related to the testing and improving of GCMs.

Secondly, the common methods of validation of GCMs as laid out in the IPCC reports is generally well-suited to mimicking the current climate state but is ill-suited to guaranteeing accurate and precise predictions of future climate states. The IPCC 1995 report opens its discussion of model evaluation by stating that it takes the approach of Oreskes et al. (1994) which endorses the matching of known climate states between model and data. Very often this is accomplished by adjustment of parameterizations of sub-gridscale processes such as cloud microphysics, air-sea interactions, turbulence, etc. When this is done, it is difficult to know whether the change is physically based, especially since the changes are frequently implemented to guarantee overall stability of the model (see discussion on flux adjustment in the IPCC 1995 report). If the physics of the model is not reasonable, then predictions of future changes in the climate state cannot be reasonably trusted. Leith (1975) pointed out the connection between second-moments in the climate system and forced evolution of the climate system. By implication, predictions of climate change by GCMs are more trustworthy if second-moments in GCM variables and observations are in good agreement. Here we outline Leith's note, some follow-up research, and how occultation data of COSMIC is relevant to Leith's idea.

2   Climate signal detection

To what degree is the warming of the last century caused by increasing concentrations of anthropogenic greenhouse gases? Could the warming possibly be a natural fluctuation of the climate system instead? In order to answer these questions, Barnett and Schlesinger (1987) suggested

that the spatial pattern of the warming ought to distinguish it from patterns of natural variability. This technique is known generally as "fingerprinting," a signal having a unique fingerprint which would distinguish it from other suspected signals. At first the fingerprinting technique was applied as a simple correlation between the expected signal and the data to be used (c.f. Santer et al. 1995). Whenever such a correlation is done, it is always expected that even in the absence of a signal-to-be-detected that some "noise" leaks through. This noise is the result of the "natural variability" of the climate, and hence research into the problem of detecting a signal came to be answered in terms of signal-to-noise ratios (Hasselmann 1979). Subsequently others used conventional signal processing methods to show that a kind of filter could be devised which would minimize the amount of noise that would "leak" into a detection—a kind of optimization of the correlation technique (Bell 1982; Bell 1986; North et al. 1995). Finally, it was shown that several different signals can be detected simultaneously in an optimal fashion (Hasselmann 1993). Here we will call this optimal multi-pattern regression.

The main premise in climate signal detection studies is that the pattern of the signal as it would appear in the data can be adequately described by a GCM (or simpler model). Although the initial aim was to detect global warming and attribute it to an enhanced greenhouse effect, in reality any "externally" forced signal can be sought in signal detection. For instance, naturally occuring phenomena, such as a change in solar insolation or suspended ash from a volcanic eruption, can be viewed as external forcings. Even phenomena internal to a climate system can be treated as external forcings, such as a sea-surface temperature warming associated with an El Niño.

2.1   Correlation studies

Assume we have a data set $\mathbf{d}$. This data set can be comprised of any atmospheric or oceanic variable of interest or any combination thereof, such as coefficients of a Fourier transform or of a spherical harmonic expansion. In most studies, the data is grouped according to time, $\mathbf{d}(t)$. It is important that the data be relative to a mean. We wish to find whether a signal is present in that data.

In all detection studies, we assume that the pattern of the signal $\mathbf{s}$ can be modeled by a computer to within a scaling factor α. Such a pattern is computed by subtracting a control run of a GCM with the desired external forcing absent from another run of the same GCM with the external forcing present. The data can then be modeled by

$$\mathbf{d} = \alpha\,\mathbf{s} + \mathbf{n}, \tag{1}$$

and in the case of a timeseries of data, the data is modeled by

$$\mathbf{d}(t) = \alpha(t)\,\mathbf{s} + \mathbf{n}(t). \tag{2}$$

For now we will consider that $\mathbf{s}$ is a strictly spatial pattern and thus the scaling factor α evolves in time. In both cases $\mathbf{n}$ represents the noise of natural climate variability as it would appear in the data. It is assumed that a mean has been subtracted from the data. In applying the correlation, the best estimate of the signal amplitude $\tilde{\alpha}$ is found taking the inner product of the data with the signal pattern:

$$\tilde{\alpha}(t) = \frac{\mathbf{s} \cdot \mathbf{d}(t)}{s^2} \tag{3}$$

in which $s^2 = \mathbf{s} \cdot \mathbf{s}$. The error associated with this estimate of the signal amplitude is found by computing the root-mean-square difference of $(\tilde{\alpha} - \alpha)$ assuming that the climate noise $\mathbf{n}$ is Gaussian. The error, σ_α, is given by

$$\sigma_\alpha^2 = \frac{\langle (\mathbf{s} \cdot \mathbf{n})^2 \rangle}{s^4}. \tag{4}$$

2

The notation $\langle \ldots \rangle$ indicates an ensemble average. When dealing with a timeseries of data $\mathbf{d}(t)$, the uncertainty is estimated by taking $\mathbf{n}$ as the output from the control run without the external forcing. The signal to noise ratio is then simply $\mathrm{SNR} = \tilde{\alpha}/\sigma_\alpha$ and the confidence level of detection is $\mathrm{erf}(\mathrm{SNR}/\sqrt{2})$.

The inner product can be defined in any of a variety of ways. The data can take the form of a vector of similar (or even non-similar) elements. For instance, it can be only temperatures, or a combination of temperatures and humidities. The data can also take the form of a functional wherein the inner product becomes an integral over volume. In any case, a problem arises because in this method the inner product does not necessarily remain invariant under coordinate transformation. This means that different combinations of the same data can yield very different results for the signal-to-noise ratio. In practice, it leaves undetermined how one should weight data in data-sparse regions: how does one assign the appropriate amount of weight to surface temperature measurements in oceanic regions where there is very little data?

While this method can give a fairly robust estimate of the confidence level of detection for a single pronounced signal, it is inappropriate should the signal be weak or other strong signals be present. The former is a concern with some global warming detection work in the literature (Santer et al. 1996; Tett et al. 1996). In that work, the authors assembled a signal shape using a linear combination of the effects of enhanced greenhouse warming, cooling by sulfate aerosols, and stratospheric cooling by ozone loss to best match the signal seen in the data. Their results are that they find the signal with a high level of confidence, but in a way this is circular reasoning. A better way to tackle the problem must be with a signal detection method which accounts for multiple signals.

2.2 Multi-pattern regression

We can adapt the correlation method above to handle the problem of multiple signals. The model becomes

$$\mathbf{d}(t) = \sum_{i=1}^{n} \alpha_i(t)\, \mathbf{s}_i + \mathbf{n}(t). \tag{5}$$

It is assumed that there are n signals present, each with its own signal pattern $\mathbf{s}_i$ and amplitude α_i which evolves in time. One can arrive at a best estimate of the signal amplitudes as a function of time by a least-χ^2 method. One defines χ^2 by

$$\chi^2(t) = \left(\mathbf{d}(t) - \sum_{i=1}^{n} \alpha_i(t)\, \mathbf{s}_i\right)^2 \tag{6}$$

and minimize it by varying each α_i at each time interval. The result is

$$\tilde{\alpha}(t) = \mathbf{G}^{-1}\, \mathbf{p}(t) \tag{7}$$

in which the elements of the matrix $\mathbf{G}$ are given by $G_{ij} = \mathbf{s}_i \cdot \mathbf{s}_j$ and the elements of the vector $\mathbf{p}(t)$ are $p_i(t) = \mathbf{s}_i \cdot \mathbf{d}(t)$. The uncertainty is given as an uncertainty covariance matrix $\tilde{\mathbf{A}}$, the elements of which are defined as $A_{ij} \equiv \langle \delta\alpha_i \delta\alpha_j \rangle$ in which $\delta\alpha_i$ refers to a departure from the best estimate of α_i. The uncertainty covariance matrix is

$$\tilde{\mathbf{A}} = \mathbf{G}^{-1}\, \mathbf{T}\, \mathbf{G}^{-1} \tag{8}$$

in which the elements of $\mathbf{T}$ are $T_{ij} = \langle (\mathbf{s}_i \cdot \mathbf{n})(\mathbf{s}_j \cdot \mathbf{n}) \rangle$.

One can check quickly that the amplitudes of the signals can be determined uniquely: the presence of one signal theoretically cannot be confused with the presence of another. This works

because in detecting a particular signal, one is effectively correlating a part of that signal which is unique to it alone and does not correlate with any of the other signals. In reality, though, it is impossible to formulate the signal patterns $\mathbf{s}_i$ perfectly, and, through $\mathbf{G}^{-1}$, the errors in their forms will cause confusion when distinguishing between signals.

Even though the difficulty of dealing with multiple signals is resolved by this method, the problem of transformation invariance remains. By taking different combinations of the same data, one can arrive at different conclusions. The only way to avoid this difficulty is to define a metric which makes χ^2 invariant under transformation. This metric will also optimize the detection, as we will see in the next section.

Curiously, this method of multi-pattern regression has not been applied in detecting climate signals. For studies such as those by Santer et al. (1996), it would only have required the computation of the 3-by-3 matrix $\mathbf{G}$.

2.3   Optimization

North et al. (1995) derived a set of coefficients to be used in correlation which minimizes the amount of noise that would leak into the detection using the theory of optimal filtering. The model is that of equation 1, and the noise is assumed Gaussian with $\mathbf{e}_\mu$ and λ_μ being the eigenvectors/eigenfunctions (empirical orthogonal functions—EOFs) and eigenvalues of the noise covariance matrix $\langle \mathbf{nn}^T \rangle$. Their result is

$$\tilde{\alpha} = \gamma^{-2}\, p \tag{9}$$

where

$$\gamma^2 = \sum_\mu \frac{(\mathbf{e}_\mu \cdot \mathbf{s})^2}{\lambda_\mu} \tag{10}$$

$$p = \sum_\mu \frac{(\mathbf{e}_\mu \cdot \mathbf{s})\,(\mathbf{e}_\mu \cdot \mathbf{d})}{\lambda_\mu}. \tag{11}$$

The sum over μ is arbitrary: any subset of the EOFs can be chosen. The error in the estimate of $\tilde{\alpha}$ is just

$$\sigma_\alpha^2 = \gamma^{-2}. \tag{12}$$

This technique resolves the problem of transformation invariance: the result only depends on the type of data used and not on a particular linear combination of the data. Also, the data (and signal pattern and EOFs) can be spatial-temporal in nature, functional or vectorial in form. For instance, if surface temperature measurements are used, the result for $\tilde{\alpha}$ would not depend on how oceanic-based stations are weighted in comparison to land-based stations in the composition of the data $\mathbf{d}$.

This optimization can be generalized to multi-pattern regression as shown in Hasselmann (1997). In this case, the best estimates for the signal detection amplitudes α are given by

$$\tilde{\alpha} = \mathbf{\Gamma}^{-1}\, \mathbf{\Pi} \tag{13}$$

where the elements of the matrix $\mathbf{\Gamma}$ and the vector $\mathbf{\Pi}$ are

$$\Gamma_{ij} = \sum_\mu \frac{(\mathbf{e}_\mu \cdot \mathbf{s}_i)\,(\mathbf{e}_\mu \cdot \mathbf{s}_j)}{\lambda_\mu} \tag{14}$$

$$\Pi_i = \sum_\mu \frac{(\mathbf{e}_\mu \cdot \mathbf{s}_i)\,(\mathbf{e}_\mu \cdot \mathbf{d})}{\lambda_\mu}. \tag{15}$$

The uncertainty covariance in the signals' amplitudes is

$$\tilde{\mathbf{A}} = \mathbf{\Gamma}^{-1}. \tag{16}$$

As in the multi-pattern regression described earlier, this method also is able to perfectly distinguish between signals with different patterns. As before, if the patterns are described imperfectly, this technique will confuse the presence of one signal for another. This last set of equations is state-of-the-art in climate signal detection. It is referred to as optimal multi-pattern regression.

The form presented here can be read in either functional form or vector form, depending on how the inner product is defined. Because ultimately this problem is applied numerically, it is most appropriate to think of the quantities $\mathbf{d}$, $\mathbf{s}_i$, and $\mathbf{e}_\mu$ as vectors. As such, the noise covariance matrix can be written as

$$\mathbf{N} = \langle \mathbf{nn}^T \rangle = \sum_\mu \mathbf{e}_\mu \mathbf{e}_\mu{}^T \, \lambda_\mu \tag{17}$$

and optimal multi-pattern regression can be found by using this noise covariance matrix as a metric in defining a new χ^2 (Hasselmann 1997; Leroy 1999):

$$\chi^2 = (\mathbf{d} - \mathbf{S}\alpha)^T \, \mathbf{N}^{-1} \, (\mathbf{d} - \mathbf{S}\alpha). \tag{18}$$

In this vector form, the i'th column of the matrix $\mathbf{S}$ is the individual pattern $\mathbf{s}_i$. Minimizing χ^2 by varying α gives

$$\tilde{\alpha} = (\mathbf{S}^T \mathbf{N}^{-1} \mathbf{S})^{-1} \, \mathbf{S}^T \mathbf{N}^{-1} \mathbf{d}, \tag{19}$$

which is the same result as given by equations 13, 14, and 15 in vector form. In practice, the inverse of the noise covariance matrix is composed of a subset of the EOFs by

$$\mathbf{N}^{-1} = \sum_\mu \mathbf{e}_\mu \mathbf{e}_\mu{}^T \, \lambda_\mu^{-1}. \tag{20}$$

In general, those EOFs with smaller associated variances λ_μ are omitted because the amount of information available to compute the noise covariance matrix $\mathbf{N}$ does not justify small variance modes (c.f. Hegerl et al. 1997). (Notice that $\mathbf{NN}^{-1}$ no longer reduces to an identity matrix but to $\sum_\mu \mathbf{e}_\mu \mathbf{e}_\mu{}^T$.)

2.4 Application of COSMIC data

Most of the global warming detection studies to date have utilized a ~ 120-year record of temperatures from meteorological stations distributed worldwide or a ~ 50-year record of radiosonde soundings. Naturally, these in situ records do not give satisfactory coverage over the oceans, so future detection studies using systems which obtain global coverage are desireable. The one space-based system which has been utilized in global warming detection studies is the Microwave Sounding Unit (MSU) on the U.S. series of TIROS Operational Vertical Sounder (TOVS) weather satellites (Spencer and Christy 1992a; Spencer and Christy 1992b). The channel 2 radiance measured by MSU is representative of a bulk temperature of the troposphere; however, this was called into question in situations of cloud cover (Hansen et al. 1995).

COSMIC will be but the first in a series of occultation constellation experiments performed by different organizations around the world. As such, we view COSMIC as the benchmark from which other constellations of GPS receivers can be used to detect global warming.

Data from COSMIC will be well-suited to the task of detecting global warming because it is highly sensitive to tropospheric warming and is subject to only very small systematic errors.
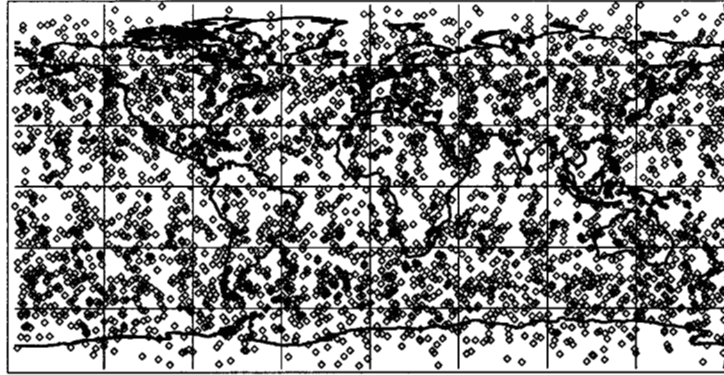
5

Figure 1: A simulated distribution of occultations by COSMIC for one day. Taken from Stevens (1998).

Leroy (1997) pointed out that GPS occultation can directly measure the geopotential heights of constant pressure surfaces from space. The measurement is dependent only on the sounding of the atmosphere *above* the layer of interest: the refractivity, which is directly proportional to density in the atmosphere above the mid-troposphere, is integrated from the top of the atmosphere downward to give pressure as a function of absolute height. By evaluating the geopotential energy as a function of the absolute height (at the horizontal position of the occultation) we deduce geopotential height as a function of pressure above the mid-troposphere. (For details on GPS occultation retrieval, see Hajj et al. 1998; Kursinski et al. 1997.) Thermodynamically, the geopotential height is a natural measure of the average temperature of the atmosphere below it:

$$h(p) = \int_p^{p_s} \frac{RT}{\mu g_0} \frac{dp'}{p'} \qquad (21)$$

where h is geopotential height, p is pressure, p_s is the surface pressure, R is the ideal gas constant, T is temperature, μ is the mean molecular mass, and g_0 is a standard value of gravitational acceleration. Thus, even though the measurement accuracy is independent of the atmospheric state below the layer of interest, the measured quantity is nonetheless a measure of the average temperature below that layer.

GPS occultation of Earth's atmosphere can only directly observe geopotential heights of constant pressure surfaces above the mid-troposphere because water vapor begins to contribute substantially to the refractivity in the lower troposphere, making the interpretation of refractivity in terms of more conventional atmospheric parameters ambiguous. While measuring the geopotential heights of pressure surfaces above the mid-troposphere is probably sufficient for global warming detection, it is valuable for reasons mentioned later to retain information on the lower atmosphere. Thus, we propose that the parameter to be retained for detection work and other work mentioned in following sections should be an *integrated refractivity* $\tilde{N}(h)$:

$$\tilde{N}(h, \mathbf{x}) \equiv a \int_h^\infty N(h', \mathbf{x}) \, dh' \qquad (22)$$

in which h is geopotential height, N is refractivity, $\mathbf{x}$ is the horizontal coordinate, and the constant $a = 0.044024$ Pa m^{-1}. Above the mid-troposphere, $\tilde{N}$ is just the atmospheric pressure.

6

In the lower troposphere, $\tilde{N}$ can be interpreted as the "weight" of refractivity above the geopotential height h which is similar to the column amount of water vapor above that height. We propose that the units of $\tilde{N}$ be called "equivalent Pascals."

We envision that the data to be used in the optimal filter will be the spherical harmonic expansion coefficients of $\tilde{N}$ at a few geopotential heights as a function of time. The coefficients will be denoted as $\tilde{N}_{lm}^{(c,s)}(h,t)$, in which l is the spherical harmonic degree, m is the order, c, s refer to a sine or cosine coefficient, h is the geopotential height, and t is the time period. The actual field is represented as

$$\tilde{N}(\phi, \theta, h, t) = \sum_{l=0}^{n} \sum_{m=0}^{l} \left(\tilde{N}_{lm}^{c}(h,t) \cos(m\phi) + \tilde{N}_{lm}^{s}(h,t) \sin(m\phi) \right) P_{lm}(\theta) \tag{23}$$

in which ϕ is longitude, θ is latitude, and $P_{lm}(\theta)$ is an associated Legendre polynomial (Abramowitz and Stegun 1972). Notice that $\tilde{N}_{lm}^{s} = 0$ when $m = 0$. For an n-degree expansion, there are $(n+1)^2$ coefficients. If we wish to retain information which describes the atmosphere on continental scales, a degree-5 expansion is required, which gives 36 coefficients.

Climate signal detection requires that data be gridded on timescales at least a month in duration. With COSMIC, roughly 110 000 occultations will be collected each month, covering the entire globe and all times-of-day (sun-fixed longitudes). Figure 1 shows a simulated distribution of soundings by COSMIC for one day. Call $\tilde{N}_i(\phi_i, \theta_i, t_i, h)$ the data from sounding i at longitude and latitude (ϕ_i, θ_i) at time t_i during a month. A set of coefficients $B_{lm,i}^{(c,s)}$ is required to convert these data to the set of spherical harmonic coefficients:

$$\tilde{N}_{lm}^{c}(h) = \sum_{i} B_{lm,i}^{c} \, \tilde{N}_i(h) \tag{24}$$

$$\tilde{N}_{lm}^{s}(h) = \sum_{i} B_{lm,i}^{s} \, \tilde{N}_i(h). \tag{25}$$

We use $\tilde{N}_i(h)$ as a shorthand for $\tilde{N}_i(\phi_i, \theta_i, t_i, h)$. There are a variety of methods to determine the coefficients $B_{lm,i}^{(c,s)}$. Since the spherical harmonic coefficients are really intended to be a measurement of the atmospheric mean over a month, we would prefer that the conversion coefficients be chosen so that the differences between the spherical harmonic coefficients determined by equations 24 and 25 and the *true* coefficients for the month are as small as possible. The implication is that an optimal method should be used to determine the conversion coefficients $B_{lm,i}^{(c,s)}$; however, the extremely large number of data involved makes the determination of the conversion coefficients prohibitive. The section on Bayesian interpolation in Leroy (1997) shows one alternative method of deriving such conversion coefficients.

The difference between the calculated spherical harmonic coefficients and the true spherical harmonic coefficients which describe the mean state of the atmosphere for the month is called "sampling" error. This error arises because the data does not completely sample the atmosphere in time and space. In between the data locations, spatially and temporally, there are maxima and minima in the integrated refractivity field which are missed by the occultation data. Albeit the variability which contributes to the error has short time and spatial scales, this variability is associated with synoptical scale weather systems. Li and North (1998) shows how to estimate the sampling error for a conversion filter such as that given by equations 24 and 25. The result is an error covariance matrix in the elements $\tilde{N}_{lm}^{(c,s)}(h)$, which should be added to the climate variability covariance matrix $\mathbf{N}$ for climate signal detection. This will alter the EOFs, generally increasing the eigenvalues λ_μ. An estimate of the sampling error for data distributed like that expected from COSMIC has not yet been done.

In addition to noise due to natural climate variability and sampling error, one must also consider systematic error. This error arises because no observing system measures atmospheric

variables perfectly. In this case, a vertical profile of refractivity deduced from an occultation, computed by an Abel transform, does not recreate the actual atmospheric profile precisely primarily because the horizontal resolution of the occultation, about 200 km, is too large to capture spatial water vapor variability in the lower troposphere. While this particular error will average to zero with 110 000 monthly occultations, there are others which are not expected to average out. These are systematic. The dominant contributor to the systematic error is caused by the ionosphere: solar-cycle variability induces variability in ionospheric activity which in turn induces error in refractivity retrievals, especially in the stratosphere. This error translates to a 100-m error in geopotential height at 40-km altitude falling exponentially to a 1-m error in the mid-troposphere. While some of this error should cancel by averaging over a month, some will not. Once this error is propagated into the values of $\tilde{N}_{lm}^{(c,s)}(h)$, its error covariance matrix should be added to the climate variability covariance matrix $\mathbf{N}$ just as the sampling error is. See Kursinski et al. (1997) for a discussion of errors related to GPS occultation.

2.5   Work to date

Using correlation and fingerprinting techniques, recent studies have shown that surface temperature increase over the past century can be attributed to anthropogenic greenhouse gas increases with a confidence level of $\approx 95\%$ (Hegerl et al. 1997 and references therein). This must be interpreted as there being a 5% chance that a natural fluctuation of the climate can fully account for the increase of 0.6 K seen over the past century. None of these studies considered that while a 0.6 K fluctuation of the global average surface temperature is rare, it is even rarer that this increase occur steadily—almost monotonically—over the course of 100 years. Even more recent studies consider the temporal structure of the signal as well, and confidence levels have increased to 99% and higher (Tett et al. 1999; Wigley et al. 1998). In short, no GCM has ever shown a global scale warming of 0.6 K occuring on a 100-year timescale in all the millenia of control runs, a point originally made by Stouffer et al. (1994).

Even though greenhouse-gas forcing seems to be required to explain the rise in surface temperature, nevertheless there are still unexplained features in the surface temperature record. The warming seems to occur in the time intervals 1910-1940 and 1970-present with little change in between. See figure 2, the data of which is described in Jones (1994) and Parker et al. (1995). When all data is included in optimal multi-pattern regression, the minimum χ^2 far exceeds the number of degrees of freedom, meaning that the fit is unsatisfactory. That is, there are still unexplained features in the temporal pattern of surface temperature over the last century. To this day this pattern of warming intervals and steady intervals remains unexplained although some of suggested that variations in solar forcing might be responsible.

Leroy (1999) and Stevens (1998) have performed optimal filtering studies which probe how global warming detection will be impacted by COSMIC. Even though no data has been obtained yet, these types of studies can still attempt to determine how warm the climate must become before we can unmistakably attribute that warming to increased concentrations of greenhouse gases. Leroy (1999) estimated that a 1-sigma detection time of detection is about 10 years, a much shorter time than if only surface temperatures are used. This result is most sensitive to the overall pattern of tropospheric warming and stratospheric cooling. Philosophically, though, a cooling stratosphere is *not* evidence that the troposphere is warming because of increased greenhouse gas concentrations. Stevens (1998) appropriately separated the tropospheric and the stratospheric patterns in determining how much each contributes to global warming detection. He found that using the thickness of the troposphere (geopotential height of the ~ 100-hPa surface) adds very little information over surface temperature measurements. This is not surprising, primarily because it is expected that temperature fluctuations are strongly coupled throughout the troposphere.

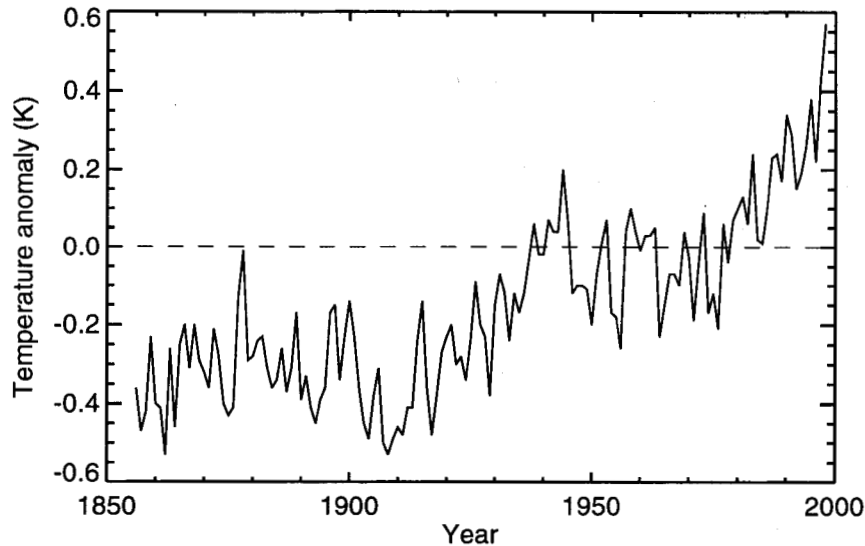Both Leroy (1999) and Stevens (1998) have demonstrated a commonly occuring problem in

Figure 2: The global average surface temperature from 1856 through 1998 as distributed by the IPCC. The data is available on the internet: http://www.cru.uea.ac.uk/cru/data/temperat.htm.

optimal detection of climate signals. It was mentioned previously that any subset of EOFs and their eigenvalues can be used in optimal detection. For example, Leroy (1999) estimated the signal pattern $\tilde{N}(h)$ for global warming in the Indian Ocean region (c.f. figure 3). In figure 4 he shows the spectrum of eigenvalues λ_μ and the square-projections $(e_\mu \cdot s)^2$. Notice that while the spectrum of λ_μ cascades strongly, the square-projections do not. By inspection of equation 10, it seems that a nearly infinite signal-to-noise ratio can be obtained for detection. This should not be true. First, as mentioned previously, there is not enough information in the control run to justify the smallest eigenvalues λ_μ. Second, the higher order EOFs tend to represent the finer, smaller-scale features of the signal pattern, details in which one cannot have much confidence. Therefore, a subjective truncation procedure must be undertaken wherein some modes are retained and the others eliminated.

Different researchers take different approaches to this truncation (c.f. Leroy 1999; Hegerl et al. 1996; Hegerl et al. 1997; North and Stevens 1998; Stevens 1998, etc.). Perhaps the most satisfying method is to use only those modes which pass a test of statistical significance before being used to parameterize the climate natural variability (Tett et al. 1999).

Most of the preparatory work required to use COSMIC occultation data for signal detection is in estimating the natural climate variability. Stevens (1998) uses CCM3 to estimate the variability of geopotential heights in the tropics, which is directly related to integrated refractivity above the mid-troposphere. Other models should be used as well, for the sake of comparison. In addition, additional information may be obtained from tracking the annual cycle. The EOFs in the annual cycle are "cyclostationary" modes (Kim et al. 1996).

In passing, it is worth noting that the fingerprinting approach to detecting climate change, premised on an approximate linearity of the climate system, might prove outdated should the view of Corti et al. (1999) be affirmed. They surmise that instead of the chaotic climate system react linearly to an external forcing that instead it visits its preferred states (attractors) with steadily changing frequency. It is not clear how fingerprinting or regression can be applied
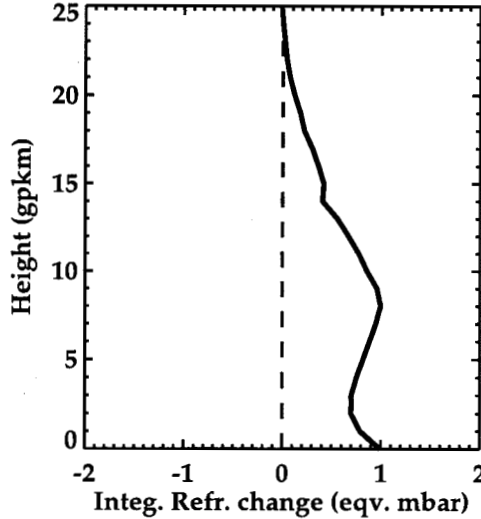
Figure 3: The signal pattern of greenhouse-gas induced warming in integrated refractivity vs. geopotential height. It is the difference in integrated refractivity profiles computed with 396 ppmv carbon dioxide and 330 ppmv using a one-dimensional radiative-convective model.

should such a mechanism be at work.

3   Detection as model testing

While climate signal detection was originally motivated by the need to detect and attribute climate change to human activity, some scientists are now using this technique to test climate models in several different ways. For example, Forest et al. (1997) shows how optimal signal detection can be used to estimate parameters of simplistic climate models. This technique can be especially appropriate for forcings on shorter timescales than global warming, such as local oceanic warmings, volcanic eruptions, etc. Also, when signal detection as presented in the previous section is formulated in Bayesian terms, it can be shown that by itself it tests climate models.

3.1   Bayesian priors and climate model predictions

In statistical language, optimal detection is finding the conditional probability of signal amplitudes given a data set. This is written as $P(\alpha|\mathbf{d})$. It is Gaussian in α with a centroid at $\tilde{\alpha}$ and an uncertainty covariance matrix of $\tilde{\mathbf{A}}$. In statistics this is written as $\alpha \sim \mathcal{N}(\tilde{\alpha}, \tilde{\mathbf{A}})$, which means

$$P(\alpha|\mathbf{d}) = (2\pi)^{-n_s/2} |\tilde{\mathbf{A}}|^{-1/2} \exp\left(-\frac{1}{2}(\alpha - \tilde{\alpha})^T \tilde{\mathbf{A}}^{-1}(\alpha - \tilde{\alpha})\right) \tag{26}$$

where n_s signals are being detected (and is the dimension of α) and $|\ldots|$ is a determinant. Recall that $\tilde{\alpha}$ is computed using the data $\mathbf{d}$ (c.f. equation 19). Implicit in finding this probability distribution for α is that the shapes of the signals, as prescribed by $\mathbf{S}$ (or $\mathbf{s}_i$), are known better a priori than the overall amplitudes of the signals α, since the signal amplitudes are treated as unknowns. In fact, subjectively we probably have a better idea of the overall amplitudes of
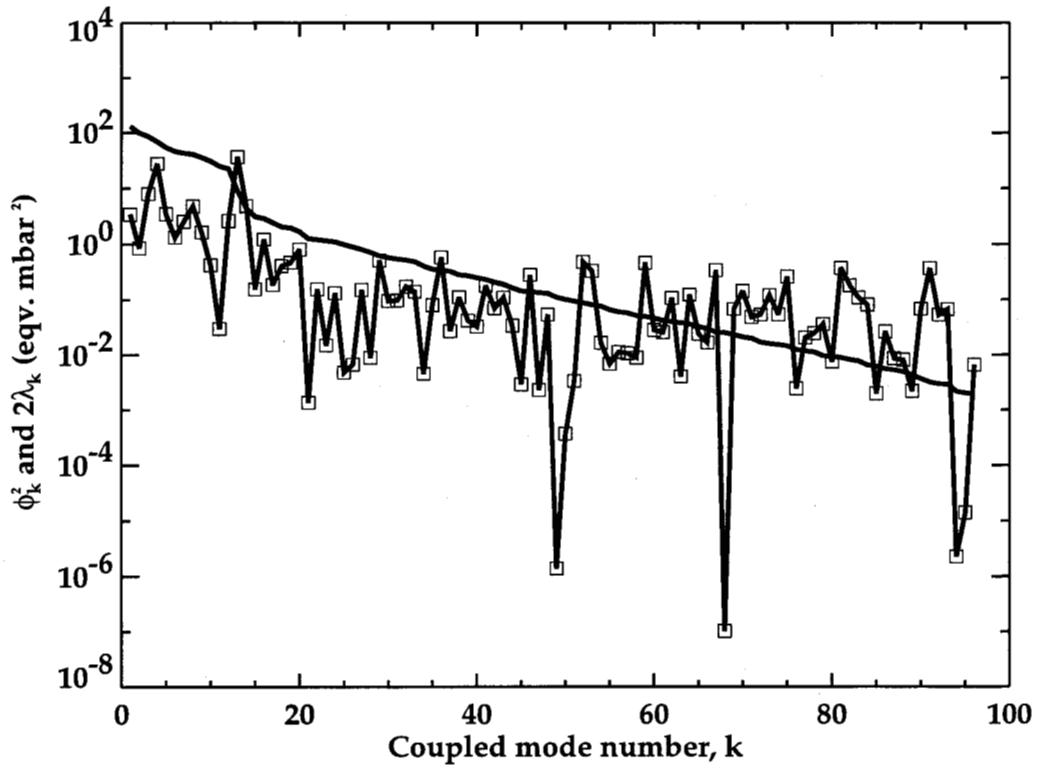
Figure 4: Eigenvalues λ_k and squared-projections $\phi_k^2 = (e_k \cdot s)^2$ for detecting the signal in figure 3 when all months-of-the-year are considered. The eigenvalues are represented by the monotonically decreasing solid line, and the squared-projections are represented by the line connecting the square points. The variability was computed using the first 100 years from the 300-yr control run of the Climate System Model (CSM) of the National Center for Atmospheric Research (NCAR) of the U.S.

signals than we do of their spatial scales at high resolution. Therefore, it is only appropriate that we account for our a priori knowledge of signal amplitudes.

In Bayesian statistics, equation 26 is only half of the problem of inference: the other half is the "prior" in the signal amplitudes, since the posterior is a probability density in the signal amplitudes. The prior is a probability distribution in signal amplitudes which depends on the theory, or model, that is used to establish the probability distribution. It is written as $p(\alpha|M)$ in which M denotes that model M is used to determine the prior probability distribution. Bayes's theorem states that

$$P(\alpha|\mathbf{d}, M) \propto P(\mathbf{d}|\alpha) \, p(\alpha|M) \qquad (27)$$

wherein $P(\alpha|\mathbf{d}, M)$ is the "posterior" (posterior probability distribution of α), $P(\mathbf{d}|\alpha)$ is the "likelihood" (likelihood of the data), and $p(\alpha|M)$ is the "prior" (a priori knowledge of α given model M). The normalization constant for the posterior $P(\mathbf{d}|M)$ is given by integrating over α:

$$P(\mathbf{d}|M) = \int P(\mathbf{d}|\alpha) \, p(\alpha|M) \, d\alpha \qquad (28)$$

The Bayesian likelihood function $P(\mathbf{d}|\alpha)$ is directly proportional to the posterior probability $P(\alpha|\mathbf{d})$ given by more traditional statistics in equation 26.

It is the prior $p(\alpha|M)$ which is uniquely Bayesian. It must be interpreted as a prior expectation of what the signal amplitudes should be before any data is considered, effectively making it subjective. Those expectations derive from the best theories of the climate system available: GCMs.

At this point it is worth noting how global warming predictions are currently done. Usually, one works with a single model M. First, a long control run is done with that model to establish how the climate system would behave without any external forcing. A climatic mean state is determined. Then the model is run with a particular external forcing imposed. The difference between the first run and the second run is the signal. Usually only the overall spatio-temporal form of the signal is retained, allowing a rescaling of the overall amplitude (as was done in the first part of this paper). In reality, we also have a prediction for the signal amplitude as well.

In order to establish a proper prior the model also needs to assign uncertainties to its predicted α. Most work assumes that the uncertainty derives from the climate variability as generated by the model; however, the true uncertainty derives from the uncertainties of the model. For instance, whereas the equations of motion incorporated into the atmospheric and oceanic models are precise, some error is incurred by implementing them in finite difference form. Most important, though, are the uncertainties associated with the parameters of the model, those parameters being necessary to describe processes in the ocean-atmosphere-biosphere-cryosphere system which occur on sub-gridscales. Examples of such parameters would be associated with schemes of moist convection, cloud formation/properties, surface run-off, etc. The uncertainties in these parameters are subjective, but they can be estimated reasonably. Preferably, then, the uncertainty in the climate signal prediction should be estimated by running the model, with the external forcing included, several times, each time with different but reasonable values of internal parameters. This is computationally extremely expensive, but alternative approaches to such a calculation have been proposed (Webster and Sokolov 1997).

Because the likelihood function is generally much more sharply peaked than the prior (otherwise there was no reason to obtain the data to begin with), the prior will have little effect on the posterior probably distribution. If more than one model is used to generate a prior, however, the prior becomes very useful in a *second* level of inference. If a relative prior probability for model j is $p(M_j)$, then the posterior relative probability of that model in light of the data $\mathbf{d}$ is $p(M_j|\mathbf{d})$:

$$p(M_j|\mathbf{d}) \propto P(\mathbf{d}|M_j) \, p(M_j) \qquad (29)$$

wherein the function $P(\mathbf{d}|M_j)$ is evaluated using equation 28. If we assume that the prior distribution in $\boldsymbol{\alpha}$—a model's prediction—is $\boldsymbol{\alpha} \sim \mathcal{N}(\boldsymbol{\alpha}_p, \mathbf{A}_p)$, then the conditional probability of the data on the model is

$$P(\mathbf{d}|M) = (2\pi)^{-n_s/2} \Big(\frac{|\mathbf{A}_m|}{|\mathbf{A}_p||\tilde{\mathbf{A}}|} \Big)^{1/2} \exp\big(-\frac{\mathcal{A}}{2}\big) \tag{30}$$

where the "accuracy" $\mathcal{A}$, the posterior most probable signal amplitudes $\boldsymbol{\alpha}_m$, and the posterior amplitude uncertainty covariance matrix are given by

$$\mathcal{A} \;\; = (\mathbf{d} - \mathbf{S}\boldsymbol{\alpha}_m)^T \mathbf{N}^{-1}(\mathbf{d} - \mathbf{S}\boldsymbol{\alpha}_m) + (\boldsymbol{\alpha}_m - \boldsymbol{\alpha}_p)^T \mathbf{A}_p^{-1}(\boldsymbol{\alpha}_m - \boldsymbol{\alpha}_p) \tag{31}$$

$$\boldsymbol{\alpha}_m = (\tilde{\mathbf{A}}^{-1} + \mathbf{A}_p^{-1})^{-1} (\tilde{\mathbf{A}}^{-1}\tilde{\boldsymbol{\alpha}} + \mathbf{A}_p^{-1}\boldsymbol{\alpha}_p) \tag{32}$$

$$\mathbf{A}_m = (\tilde{\mathbf{A}}^{-1} + \mathbf{A}_p^{-1})^{-1}. \tag{33}$$

Throughout these equations the subscript p refers to the prior probability distribution, the subscript m refers to the posterior probability distribution, and the tilde refers to the data likelihood function. The accuracy $\mathcal{A}$ given by equation 31 can be thought of as the accuracy of the prediction. The first term on the right is the disagreement between the data and the most likely fit described by $\boldsymbol{\alpha}_m$, and the second term on the right is the disagreement between the outcome and the prior prediction. The higher the accuracy $\mathcal{A}$ the less accurate the result.

The conditional probability of the data on the model gives a relative probability of one model to another model. Equation 28 indicates that the less accurate the result, the less relatively probable the model is in light of the data. This is obvious from experience. What is new with a Bayesian formulation is that the more precise the prediction (re: the smaller the prior uncertainty covariance matrix $|\mathbf{A}_p|$) the more relatively probable the model is in light of the data. A model is more useful scientifically if it can give more precise predictions than another model while maintaining good accuracy. What is understood intuitively about precision is made quantitative with Bayesian inference. In this way signal detection can be shown to be a special case of model testing (Leroy 1998).

3.2 Testing models by detecting signals with COSMIC

While detecting warming associated with increased greenhouse-gases has been discussed at length, the duration of the resulting climate change makes it unsuitable for use in testing and improving GCMs. There are other external forcings which are better suited, e.g. a sea-surface temperature warming associated with an El-Niño even or a volcanic eruption. For an El Niño forcing, we expect the dominant model response to be a change in the global pattern of precipitation and in patterns of tropical convection. The relevant model parameterizations would be those associated with surface-atmosphere interaction, with moist convection, and with cloud formation and evolution. For volcanic forcing, we expect the dominant model response to be regional cooling of the troposphere. The pattern of such a cooling bears upon radiative transfer parameterizations and circulation/transpose in the stratosphere.

Differences between modeled and observed responses to external forcings will not only be in the amplitude but also in the overall pattern. Since the pattern is strictly defined in fingerprinting, it seems that the method given above may not be well-suited to improving climate models. The only improvement to a model will be an adjustment of its parameters, something that is best done with a climate adjoint. Such an adjoint would be similar to those used in weather forecasting (c.f. Hall and Cacuci 1982 and Errico 1997). We expect the prime difference between a climate adjoint and a weather adjoint to be that the parameters of the model are adjusted instead of the initial atmospheric state. The development of a climate adjoint is a very new field and everything remains to be done.

4 Testing models using second-moment statistics

Climate signal detection clearly relies heavily on climate models: first (and foremost) to prescribe the natural variability of the climate in the variables of interest, and also to specify the patterns of the forced signals for which we are searching. In order to get a sense of how reliably a climate model can describe the climate variability, it is necessary to use another method to test the climate model. Leith (1975) proposed a method which is ideally suited to this task. His idea is based on the application of the fluctuation-dissipation theorem of thermodynamics to the ocean-atmosphere system.

The outcome of his note is that the quality of a model is best tested by comparing second-moment statistics of its variables. Most climate model validation efforts involve a comparison of model averages and observed means. In the event of a mismatch it is simple enough to tune the model to match model output and observations regardless of the physical reality of the climate system. By paying attention to second-moment statistics of climate variables, though, one is more likely to realistically simulate the actual physics of the climate system, even if long-term averages may not necessarily match reality. Any mismatch of averages then represents the absence of a physical process in the model rather than an implementation of incorrect physics. Hence, we regard a model which is validated by a comparison of second-moment statistics as a more useful tool for prediction than one whose first-order moments are compared.

We first describe Leith's note qualitatively. Then we describe the work of some authors who have used this idea to test climate models, including some work which uses second-moment statistics of observed field to estimate internal feedbacks in the climate system.

4.1 Leith's note

Leith (1975) sought to show that there is a link between the second-order moments of a climate system and its sensitivity to external forcings. He applied the fluctuation dissipation theorem, even though the theorem holds only for Liouville sytems. The climate system is not Liouville, but subsequent work demonstrated that the theorem remains applicable to the climate system to a large degree (North, Bell, and Hardin 1993). An heuristic explanation of Leith's work is presented here.

Consider the climate system as characterized by a large set of variables u_α. We wish to find how those variables respond to an external forcing of one of those variables. The response as a function of time can be written in terms of a Green's function:

$$\delta u_\alpha(t) = \int_{-\infty}^t g_{\alpha\beta}(t - t')\, \delta f_\beta(t')\, dt' \tag{34}$$

in which δu_α is the response of variable u_α to an external forcing δf_β in variable u_β and $g_{\alpha\beta}(\tau)$ is the response function.

Again, assume the system is Liouville (energy-conserving) with its variables fluctuating randomly, each fluctuation of a mode leading to responses in all the other modes through weak interactions. All of the modes are in a type of statistical equilibrium, like an equipartition of energy $(k_B T)$ in thermodynamics. One can then think of a flutuation δf_β leading to a fluctuation δu_α through a weak coupling, and one finds that

$$g'_{\gamma\sigma}(\tau) = \frac{\langle \delta u'_\gamma(t)\, \delta u'_\sigma(t - \tau) \rangle}{\langle \delta u'_\sigma\, \delta u'_\sigma \rangle} \tag{35}$$

in which $\delta u'_\gamma$ is the amplitude of the fluctuation of a mode γ, each mode being a linear combination of the variables u_α and is orthogonal to every other mode. The denominator of equation 35 can be thought of as a forcing by the random fluctuations of model σ and the numerator as
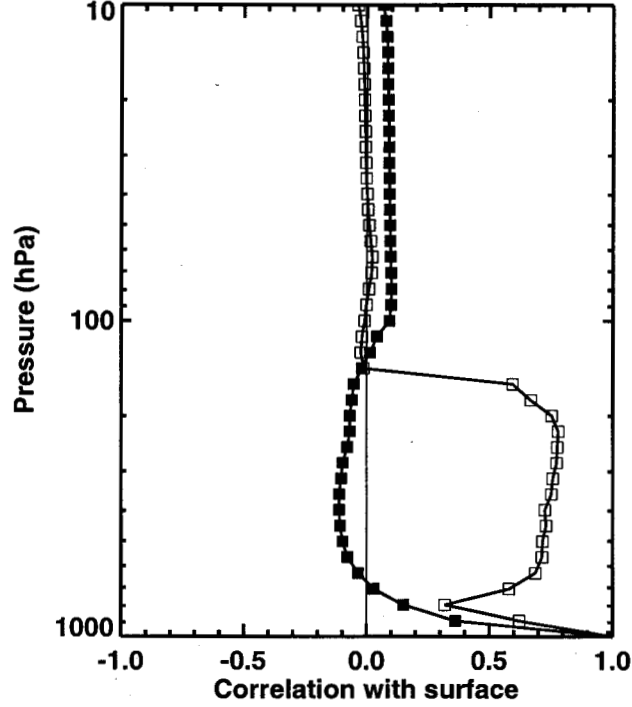
14

Figure 5: The monthly correlation of the specific humidity at the surface with upper air specific humidity over the Sahara desert (filled squares) and over the Indian Ocean (open squares). Computed from the output of the 300-year control of the CCM3.

the response to that forcing. When rewritten in terms of the variables of the system, equation 35 becomes

$$\mathbf{g}(\tau) = \mathbf{U}(\tau) \ \mathbf{U}^{-1}(0) \tag{36}$$

in which $\mathbf{U}(\tau)$ is the time-lagged covariance matrix of the atmospheric variables.

Because the time-lagged covariances can be computed from observations, it seems that the climate sensitivity can be derived from observations alone. Nevertheless, it is impractical to compute the climate sensitivity this way because of the number of variables involved. On the other hand, it is clear that it is the second-moments of the climate system which determine its sensitivity, and thus it is advised that the second-order moments of a predictive model agree well with reality.

4.2   Linear models as a means of testing GCMs

Recently, Polyak et al. (Polyak 1996; Polyak and North 1997a; Polyak and North 1997b) have described a method which is linearized version of the Leith problem. They write a linear model of the natural variability of the zonally averaged surface temperature field T_i for different latitude belts i. :

$$\frac{dT_i}{dt} + \sum_j A_{ij} T_j = \epsilon_i(t) \tag{37}$$

One can find the linear coefficients A_{ij} by optimal statistical methods for both data and control runs of models and compare. Since the records are short one must assess statistical significance of the differences. It is reckoned that the values of the coefficients are very important since in principal they are the basis for feedbacks in the system if it is perturbed by time independent (static) external influences. Such a static perturbation (e.g., doubling CO_2) is a standard measure of the sensitivity of a model. For example, let there be an external perturbation ΔQ_i, then

$$\Delta T_i \propto \sum_j \{A^{-1}\}_{ij} \Delta Q_j \tag{38}$$

The coefficients are estimated from data streams from data and generated by a GCM. They are sensitively dependent upon the lagged space-time covariances in these two streams. If the model does not get these coefficients right it is not likely to get the sensitivity right. Hence, we feel that getting these coefficients right is a necessary condition for model success in simulating climate change. While the cited papers open the door to studying this problem, much has to be done to improve our understanding of the process. First of all the studies mentioned do not take into account the fact signals are contained in the data stream. These induce correlation structure that is not in the unforced control runs. One must either add the signals to the model runs or remove some approximation to them from the data stream. This is work yet to be done.

Because of the connection between model sensitivity and the lagged space-time statistics, one would like to compare such second moment statistics for fields other than the surface temperature. Most observations taken away from the surface are not very reliable because of poor sampling and the many biases associated with such measurements. The data expected to come from the COSMIC program should provide a unique source of second moment data for comparison with GCM output. This will pose a dramatic new data set with which to confront the GCM community. It will be a data set which bears directly on the ability of the GCM to respond to external perturbations.

While any observeable related to the climate system can be used in second-moment statistics studies, some data is better at revealing specific processes than others. Occultations using COSMIC we anticipate to reveal a great deal about processes related to the water vapor cycle in the lower troposphere. Occultation by GNSS is sensitive to lower tropospheric water vapor like no other remote sensing platform.

A typical second-moment of interest would be the correlation between refractivity at the surface vs. refractivity at altitude. It is hypothesized by Lindzen (1995) (and references therein) that humidity in the upper troposphere is anticorrelated with humidity in the lower troposphere, making a negative feedback to increased shortwave forcing at the surface. A second-moment statistic related to such a process would be the correlation between surface specific humidity and upper tropospheric humidity. In figure 5 we show such a correlation for two regions as exists in the 300-year control run of CCM3 (Boville and Gent 1998). While specific humidity is not directly retrievable from atmospheric refractivity, its contribution to refractivity variance in the lower troposphere is about 100 times larger than the contribution to refractivity of the dry air. Thus, vertical correlations of refractivity from occultation data should inform us mostly on water vapor physics and transport in GCMs.

Another important second-moment would be spatial correlation of upper tropospheric and stratospheric integrated refractivities in polar latitudes. The arctic oscillation (Thompson and Wallace 1998; Shindell et al. 1999) ought to be easily detected in such hemisphere-wide correlations. Furthermore, COSMIC offers the possibility of finding a similar oscillation in the southern hemisphere where upper air sounding is much less frequent than in the northern hemisphere.

16

References

Barnett, T. P. and M. E. Schlesinger (1987). Detecting changes in global climate induced by greenhouse gases. *J. Geophys. Res. 92*, 14772–14780.

Bell, T. L. (1982). Optimal weighting of data to detect climatic change: Application to the carbon dioxide problem. *J. Geophys. Res. 87*, 11161–11170.

Bell, T. L. (1986). Theory of optimal weighting of data to detect climate change. *J. Atmos. Sci. 43*, 1694–1710.

Boville, B. A. and P. R. Gent (1998). The NCAR climate system model, version one. *J. Climate 11*, 1115–1130.

Corti, S., F. Molteni, and T. N. Palmer (1999). Signature of recent climate change in frequencies of natural atmospheric circulation regimes. *Nature 398*, 799–802.

Errico, R. M. (1997). What is an adjoint model? *Bull. Am. Meteor. Soc. 79*, 2577–2591.

Forest, C., M. R. Allen, and A. Sokolov (1997, Fall). Constraining climate model parameters using climate observations. In *Transactions, AGU*, San Francisco, California. American Geophysical Union.

Hajj, G. A., E. R. Kursinski, L. J. Romans, W. I. Bertiger, and S. S. Leroy (1998). A technical description of atmospheric sensing via GPS occultation. Submitted to J. Atmos. Ocean. Tech.

Hall, M. C. G. and D. G. Cacuci (1982). Sensitivity analysis of a radiative-convective model by the adjoint method. *J. Atmos. Sci. 39*, 2038–2050.

Hansen, J., H. Wilson, M. Sato, R. Ruedy, K. Shah, and E. Hansen (1995). Satellite and surface temperature data at odds? *Climate Change 30*, 103–117.

Hasselmann, K. (1979). *On the Signal-to-Noise Problem in Atmospheric Response Studies*, pp. 251–259. Royal Meteorological Society London.

Hasselmann, K. (1993). Optimal fingerprints for the detection of time-dependent climate-change. *J. Climate 6*, 1957–1971.

Hasselmann, K. (1997). Multi-pattern fingerprint method for detection and attribution of climate-change. *Climate Dyn. 13*, 601–611.

Hegerl, G. C., K. Hasselmann, U. Cubasch, J. F. B. Mitchell, E. Roeckner, R. Voss, and J. Waszkewitz (1997). Multi-fingerprint detection and attribution analysis of greenhouse gas, greenhouse gas-plus-aerosol and solar forced climate change. *Climate Dyn. 13*, 613–634.

Hegerl, G. C., H. von Storch, K. Hasselmann, B. D. Santer, U. Cubasch, and P. D. Jones (1996). Detecting greenhouse-gas-induced climate change with an optimal fingerprint method. *J. Climate 9*, 2281–2306.

Houghton, J., G. Jenkins, and J. Ephraums (Eds.) (1991). *Climate Change. The IPCC Scientific Assessment*. Cambridge: Cambridge University Press.

Houghton, J. and L. G. Meira Filho (Eds.) (1996). *Climate Change 1995. The IPCC Second Scientific Assessment*. Cambridge: Cambridge University Press.

Jones, P. D. (1994). Hemispheric surface air temperature variations: a reanalysis and an update to 1993. *JCLIM 7*, 1794–1802.

Kim, K.-Y., G. R. North, and J. Huang (1996). EOFs of one-dimensional cyclostationary time series: Computations, exmaples, and stochastic modeling. *J. Atmos. Sci. 53*, 1007–1017.

Kursinski, E. R., G. A. Hajj, J. T. Schofield, R. P. Linfield, and K. R. Hardy (1997). Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System. *J. Geophys. Res. 102*, 23429–23465.

Leith, C. E. (1975). Climate response and fluctuation dissipation. *J. Atmos. Sci. 32*, 2022–2026.

Leroy, S. S. (1997). Measurement of geopotential heights by GPS radio occultation. *J. Geophys. Res. 102*, 6971–6986.

Leroy, S. S. (1998). Detecting climate signals: Some bayesian aspects. *J. Climate 11*, 640–651.

Leroy, S. S. (1999). Optimal detection of global warming using temperature profiles: A methodology. *J. Climate 12*, 1185–1198.

Li, T.-H. and G. R. North (1998). Sampling errors for meteorological fields observed by low Earth-orbiting satellites. Submitted to JGR.

Lindzen, R. S. (1995). Can increasing carbon-dioxide cause climate-change? *Proc. Nat. Acad. Sci. 94*, 8335–8342.

North, G. R., R. E. Bell, and J. W. Hardin (1993). Fluctuation dissipation in a general-circulation model. *Climate Dyn. 8*, 259–264.

North, G. R., K.-Y. Kim, S. S. Shen, and J. W. Hardin (1995). Detection of forced climate signals. Part I: Filter theory. *J. Climate 8*, 401–408.

North, G. R. and M. J. Stevens (1998). Detecting climate signals in the surface-temperature record. *J. Climate 11*, 563–577.

Oreskes, N., K. Shrader-Frechette, and K. Belitz (1994). Verification, validation and confirmation of numerical models in the earth sciences. *Science 263*, 641–646.

Parker, D. E., C. K. Folland, and M. Jackson (1995). Marine surface air temperature: Observed variations and data requirements. *Climatic Change 31*, 559–600.

Polyak, I. (1996). Observed versus simulated second-moment climate statistics in GCM verification problems. *J. Atmos. Sci. 53*, 677–694.

Polyak, I. and G. North (1997a). Evaluation of the geophysical fluid dynamics laboratory general-circulation model climate variability: 1. variances and zonal time-series. *J. Geophys. Res. 102*, 1921–1929.

Polyak, I. and G. North (1997b). Evaluation of the GFDL GCM climate variability 2. stochastic modeling and latitude-temporal fields. *J. Geophys. Res. 102*, 6799–6812.

Santer, B. D., K. E. Taylor, T. M. Wigley, J. E. Penner, P. D. Jones, and U. Cubasch (1995). Towards the detection and attribution of an anthropogenic effect on climate. *Clim. Dyn. 12*, 79–100.

Santer, B. D., K. E. Taylor, T. M. L. Wigley, P. D. Jones, D. J. Karoly, J. F. B. Mitchell, A. H. Oort, J. E. Penner, V. Ramaswamy, M. Schwarzkopf, R. S. Stouffer, and S. F. B. Tett (1996). A search for human influences on the thermal structure in the atmosphere. *Nature 382*, 39–46.

Shindell, D. T., R. L. Miller, G. A. Schmidt, and L. Pandolfo (1999). Simulation of recent northern winter climate trends by greenhouse-gas forcing. *Nature 399*, 452–455.

Spencer, R. W. and J. R. Christy (1992a). Precision and radiosonde validation of satellite gridpoint temperature anomalies. part I: MSU channel 2. *J. Climate 5*, 847–857.

Spencer, R. W. and J. R. Christy (1992b). Precision and radiosonde validation of satellite gridpoint temperature anomalies. part II: A tropospheric retrieval and trends during 1979–90. *J. Climate 5*, 858–866.

Stevens, M. J. (1998). Optimal climate signal detection in four dimensions. Submitted to JGR.

Stouffer, R., S. Manabe, and K. Vinnikov (1994). Model assessment of the role of natural variability in recent global warming. *Nature 367*(6464), 634–636.

Tett, S. F. B., J. F. B. Mitchell, D. E. Parker, and M. R. Allen (1996). Human influence on the atmospheric vertical temperature structure: Detection and observations. *Science 274*, 1170–1173.

Tett, S. F. B., P. A. Stott, M. R. Allen, W. J. Ingram, and J. F. B. Mitchell (1999). Causes of twentieth century temperature change. *Nature 399*, 569–572.

Thompson, D. W. J. and J. M. Wallace (1998). The arctic oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett. 25*, 1297–1300.

Webster, M. D. and A. P. Sokolov (1997, Fall). Quantifying the uncertainty in a climate model. In *Transactions, AGU*, San Francisco, California. American Geophysical Union.

Wigley, T. M. L., R. L. Smith, and B. D. Santer (1998). Anthropogenic influence on the autocorrelation structure of hemispheric-mean temperatures. *Science 282*, 1676–1678.